SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

DTIC FILE COPY (2)

1a. REPORT SECURITY CLASSIFICATION
Unclassified

1b. RESTRICTIVE MARKINGS
N/A

AD-A212 973

3. DISTRIBUTION / AVAILABILITY OF REPORT
Unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
Final Report

5. MONITORING ORGANIZATION REPORT NUMBER(S)
AFOSR-TR- 89-1255

6a. NAME OF PERFORMING ORGANIZATION
Systems Engineering, Inc.

6b. OFFICE SYMBOL (If applicable)
3U587

7a. NAME OF MONITORING ORGANIZATION
Air Force Office of Scientific Research

6c. ADDRESS (City, State, and ZIP Code)
[illegible] Drive, Suite 620
Greenbelt, MD [illegible]

7b. ADDRESS (City, State, and ZIP Code)
Bolling AFB, DC 20332-6448

8a. NAME OF FUNDING / SPONSORING ORGANIZATION
AFOSR

8b. OFFICE SYMBOL (If applicable)
NP

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
F49620-88-C-0026

8c. ADDRESS (City, State, and ZIP Code)
Bldg 410
Bolling AFB DC 20332-6448

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO
63221C	D822	F1	

11. TITLE (Include Security Classification)
Algorithm Development for SDI Weapons System Allocation

12. PERSONAL AUTHOR(S)
Gilmer L. Blankenship; William Bennett

13a. TYPE OF REPORT
Final

13b. TIME COVERED
FROM 12-87 TO 5-89

14. DATE OF REPORT (Year, Month, Day)
89 06 01

15. PAGE COUNT
53

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
See attached.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)
See attached

DTIC
ELECTE
SEP 28 1989
S D
B

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL
~~LTC Nelson~~ LTC CROWLEY

22b. TELEPHONE (Include Area Code)
(202) 767-4939

22c. OFFICE SYMBOL
NM

DD FORM 1473, 84 MAR
83 APR edition may be used until exhausted.
All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

89 9 27 026

AFOSR-TR- 89-1255

ALGORITHM DEVELOPMENT FOR SDI WEAPONS SYSTEM ALLOCATION

Final Report

Contract No: F49620-88-C-0026

Submitted to:

Air Force Office of Scientific Research

Bolling AFB

Washington, DC

Attn: J. Crowley

Systems Engineering, Inc.[1]

7833 Walker Drive, Suite 620

Greenbelt, Maryland 20770

(301)-345-0375

[1]Now doing business as Techno-Sciences, Inc., 7833 Walker Drive, Suite 620, Greenbelt, Maryland 20770. (301)-345-0375.

ALGORITHM DEVELOPMENT FOR SDI WEAPONS SYSTEM ALLOCATION

Abstract:

While several SDI weapons systems can provide adequate defense in a one–on–one basis, a coordinated attack by several enemy missiles launched over a substantial volume will be difficult to resist without an efficient command and control system for warfare coordination. Our study of weapons allocation - coordination algorithms, is based on dynamical models for the missile/decoy systems including noise effects and uncertainties in the model parameters. Performance of the weapons targeting system may be measured in terms of the expected number of targets eliminated in a given interval (phase of operations) or the expected time to eliminate all the targets in a given region. Scheduling weapons deployment is a problem of constrained optimal stochastic scheduling and resource allocation for a system with many controls (weapons) and state variables. The selection of weapons deployment tactics is based on solution of a complex optimization problem. We have conducted an investigation of advanced modeling, stochastic control, and scheduling methodologies for aspects of the SDI weapons allocation problem - several platforms with assets of different character defending againest a diverse collection of targets. The models for such scenarios lead to stochastic scheduling problems which can not be handled by conventional analytical methods. We describe several different analytical approaches which have the potential for synthesis of effective engagement algorithms.

Key Words: Weapons allocation, stochastic sequencing and scheduling, index rules.

Contents

Executive Summary:

We have conducted an investigation of advanced modeling, stochastic control, and scheduling methodologies for aspects of the SDI weapons allocation problem – several platforms with assets of different character defending againest a diverse collection of targets. The models for such scenarios lead to stochastic scheduling problems which can not be handled by conventional analytical methods. We discuss several different analytical approaches which have the potential for synthesis of effective engagement algorithms.

Key Words: Weapons allocation, stochastic sequencing and scheduling, index rules.

1 Identification and Significance of the Problem

While several SDI weapons systems can provide adequate defense in a one–on–one basis, a coordinated attack by several enemy missiles launched over a substantial volume[2] will be difficult to resist without an efficient command and control system for warfare coordination. Battle Management systems for a region or the weapons allocation systems for individual stations in the region, require automated decision–making systems to rapidly evaluate the alternative actions and select deployment schemes compatible with tactical and strategic doctrines. A starting point for the development of such a system is the analysis of the coordination of various spatially–separated platforms/stations, each with multiple capabilities, to defend against several threats attacking simultaneously.

Coordination and contextual information is essential in tactical weapons deployment. The use of certain systems increases the visibility of the platform to a degree determined in part by its current position and attitude. Expendable weapons and countermeasures must be "rationed" over the course of an engagement. The interaction of weapons and countermeasures can be constructive or it can hinder performance, depending on use and the operational context. Both asynchronous and synchronous operating policies for resources can be useful in a given situation.

[2]In space and time.

The weapons platforms and the targets will undergo significant movements in their orbital paths during the (20 minute) duration of the mid-course phase.[3] Therefore, it is necessary to use dynamical models to describe engagements during this phase. Since there may be significant uncertainty in the measurements of target (and decoy) trajectories and profiles, it is necessary to use models which account for this uncertainty.

Accordingly, our study of weapons allocation algorithms, is based on dynamical models for the missile/decoy systems including noise effects and uncertainties in the model parameters. Performance of the weapons targeting system may be measured in terms of the expected number of targets eliminated in a given interval (phase of operations) or the expected time to eliminate all the targets in a given region. Other performance measures are possible.

In this framework scheduling weapons deployment becomes a problem of constrained optimal stochastic scheduling and resource allocation for a system with many controls (weapons) and state variables. The selection of weapons deployment tactics is based on solution of a complex optimization problem. The computational complexity of this problem (number of variables which must be computed) grows (at least) exponentially with the number of state variables in the system. Since this is a function of the number of targets, the computational problem is intractable in target rich environments.

Thus, the complexity of the SDI weapons allocation problem requires

[3] In this project we shall focus attention on this phase.

the use of algorithms incorporating not only advanced numerical techniques but also heuristic procedures and efficient knowledge representation methods to achieve performance levels approaching the announced SDI operational requirements. Within the limited setting of this Phase I project we have evaluated such techniques in the context of a systematic class of analytical models for management of engagements under uncertainty.

2 Problem Description

We consider stochastic control and scheduling formulations for certain aspects of the SDI weapons allocation problem – several platforms with assets of different character defending against a diverse collection of targets. The models for such scenarios lead to stochastic scheduling problems which can not be handled by conventional analytical methods. We discuss several different analytical approaches which have the potential for synthesis of effective engagement algorithms.

Since the weapons platforms are spatially distributed and mobile, distributed processing, communications, and decision–making capabilities enhance the reliability and survivability of the BM weapons C^2 system. While our primary effort has focused on the management of a single platform with multiple resources, we have also examined models for a multiple platform system with mobile command and operational units. We use stochastic scheduling methodologies to optimize the performance of each platform.

2.1 Distributed Layered Structure for Weapons System Management

In an engagement scenario, when several platforms deploy a variety of weapons against several threats, conflicts and constraints arise. Timing or precedence contraints are especially important. Erroneous threat type identification may substantially reduce reaction time and choice of weapons system response.

These observations suggest a layered control structure for weapons system coordination. At the lower level, each individual station requires a resource allocation algorithm capable of operating in a random environment in the presence of time–precedence constraints [20, 21, 30, 53]. There are several methodologies for such problems, including some promising recent developments [3, 2, 35, 36, 37, 41, 49, 50]. At the higher level, when several stations are involved, inaccuracies in threat identification may be anticipated and significant communication requirements arise. For example, careful timing is necessary to successfully "hand" a target from one station to another. Difficult questions regarding synchronous or asynchronous operation also arise when deployment occurs under changing network topology and variable local data bases.

In this report we focus on the activities of the lowest operational level in the hierarchy; however, the analytical tools used and developed are sufficiently general that they can be brought to bear on many aspects of other operational problems hierarchy. The abstract scheduling methodology is es-

pecially germaine.

2.2 Algorithm Development for Weapons System Allocation

The multiple station area weapons allocation coordination problem is a version of the multi-server scheduling problem. Since few optimal algorithms are known for this class of problems, we have examined a class of suboptimal strategies based on the distributed, hierarchical structure of the system. In this setup, each station is controlled by an "agent" which may be a computer. The agent executes a "local" control strategy to deploy weapons assets to engage threats in his area. Constraints on the deployment of weapons by neighboring agents assure that interference does not arise. Since a system with a single "command center" is not survivable, we assume that there are several BM systems as described above; and that they share a common data base. Since the areas of influence of different platforms may overlap, and since threats may pass through the areas controlled by several platforms, coordination of the weapons allocation is essential.

The agent's decision problem is to effectively engage threats in his area, his performance measure is a "reward" for successful engagement of a threat and a "penalty" for threats not engaged. The BM systems decision problem is to see that threats are (continuously) successfully engaged as they pass through the total area of influence of the agents under its command. Its

performance measure also includes a penalty for threats not engaged, and possibly penalties for revealing the position of friendly units. The information transmitted from agents to the BM system and vice versa will be summary status information.

The stochastic scheduling model we use to represent the decision problem modeled in this way may be solved by invoking strategies based on a *priority index rule*. The *index* is a scalar quantity associated with each weapons system. (Indicies may also be associated with the targets.) Its numerical values depend on the state of the weapon system, the threat data, and operational constraints imposed on the systems actions. To solve his "local scheduling problem," the agent computes the vector of indices for his resources (and in some cases indicies for the targets) and implements the resource with the largest index (or attacks the target with the largest index).

Thus, the state of the station's weapon system is described by a vector of priority indices. This is the (summary) information the platform communicates to superiors and other agents. In the hierarchical structure BM commanders can effectively direct the actions of station/platform agents by imposing constraints and performance bounds. The latter are essentially Lagrange multipliers, sometimes called "coordination variables," in system theory. These variables may be updated less frequently than the natural frequency of agent actions.

2.3 Data Requirements and Implementation Issues

The data base requirements for this system include:

- weapons platform states;
- data on threats; operational data on other components; and
- communication network status.

The data base is distributed throughout the BM system. An agent's access to the data is limited primarily by its communications and processing abilities. Distributed communications and processing facilities will be required in the weapons allocations subsystem to achieve the operational flexibility, effectiveness, and survivability mandated by the SDI program.

The complexity of the area weapons system C^2 problem and the large number of state variables involved prohibit the computation of exact "optimal" command and control strategies for each operational state, network configuration, and threat scenario. Effective (suboptimal) coordination of the weapons allocation system requires more than efficient computational algorithms; it requires a logical support structure, to delineate command options, likely interference effects, etc. to the BM station. Its primary function would be to keep track of precedence constraints effecting the deployment tactics of neighboring platforms, to guide the procedure of "handing a threat" from one platform to its neighbor to assure continuous engagement of the threat's system, and to manage interplatform and intersystem communications. During

this phase, we have not undertaken the development of a logic programming capability for this purpose as part of the allocation algorithm.

We shall use the framework of cooperative team theoretic solutions to "large scale" scheduling (allocation) problems as the analytical basis for the development of efficient algorithms. As we shall argue below, this framework provides a systematic basis for the construction of "suboptimal" but satisfactory tactics for weapon allocation. It also provides analytical procedures for the evaluation of performance, degrees of "optimality," and the evaluation of satisfactory solutions.

2.4 Summary

We use stochastic dynamical models to represent the interaction of the weapons platforms and the target systems during the post-boost and mid-course phases of operations included in this project. Our computational algorithms treat these models by various discretization procedures that ultimately reducing them to discrete time Markov chain systems [33].[4]

Based on the engagement models, we use a two step approach to the development of weapons allocation algorithms. First, we have developed a prototype set of algorithms using the *stochastic gradient method*. This method has proven effective in the treatment of large scale network planning

[4]Further work is recommended to enhance the models to include more realistic trajectory and orbital dynamics and better noise representation.

problems involving a number of state variables comparable to the dimensions one might expect for subsystems of the weapons – target engagement system. These algorithms are described below.

The algorithms computed by this method may be used as a baseline for the development of more representative strategies which reflect the operational structure of the SDI BM weapons allocation system. We have used stochastic scheduling models and *index rules* to derive dynamic engagement tactics for a BM system involving several weapons platforms responding to a large number of targets over an extended region of space. This class of analytical procedures and the format of the resulting algorithms is described below.

As we show, the technical problem of "coordinating" the actions of several complex weapons platforms is highly nontrivial. Conventional optimization procedures will not be effective – there is, in fact, no theory to support such a development. For "practical purposes," it is therefore appropriate to supplement the algorithms with a a "logical support system"[5] using certain AI techniques for "constraint directed scheduling."

[5]E.g., a "real time expert system."

3 Analytical Models

3.1 An Abstract Model for Engagement Dynamics

We will discuss an abstract mathematical version of this model to illustrate how the modeling problems and the design of "practical" algorithms based on such models may be developed. There are three important points which we wish to stress before describing the analysis:

First, it is not possible to compute "optimal" control laws for the weapons platforms in a realistic model of the weapons – target interaction. The computational burden grows unavoidably, exponentially with the number of state variables. Since each target (and decoy) will have a minimum of six state variables, and there may be thousands of targets and decoys, this is an insurmountable problem which cannot be solved using any conceivable computer technology.

Second, contrary to what one might expect, the computational problems associated with "local control" strategies are worse than those for the evaluation of global optimal strategies. That is, the problem of computing feedback controls (engagement tactics) which have been partitioned to respond to a subset of the states of the target population is more demanding than the optimal control/allocation problem for the system taken as a whole.

Third, if, however, the system dynamics have some special structural properties, or more precisely if one can approximate the target system model

by one with special properties, then it is possible to derive computationally feasible procedures for the evaluation of control/allocation strategies. Two such properties are *uncoupled dynamics* (target to target) and the case when the "equilibrium" probability distribution of the target states has a *product form*. The first case applies to targets which are individual missiles released from distinct launching sites. The second case may be suitable as a description of the distribution of a family of decoys released by a single missile (i.e., the initial dependency would diminish as the decoy constellation assumes its ultimate distribution).

To appreciate these points, consider the abstract stochastic control problem

$$\begin{gathered} V(s,x) = \min_{u \in U}\{E[\int_s^T e^{-\lambda t} c(x(t),u(t))dt | x(0) = x]\} \\ \dot{x} = b(x(t),u(t)) + \dot{w}(t) \qquad (1) \\ x(t) \in X \subset R^N, \quad u(t) \in U \subset R^M \end{gathered}$$

As a model for a specific SDI operational phase, the composite state vector $x(t)$ contains the state vectors of the various targets, the weapons platforms (position, velocity, attitude, etc.), and states for any auxillary processes (noise) which may be needed to complete specification of the dynamics of all the interacting systems. The control vector $u(t)$ represents the parameters of the weapons systems involved in the interaction which may be manipulated to direct the weapons to attack selected targets, e.g., platform alignment. The vector $\dot{w}(t)$ is a noise process, nominally a "white noise." The function b describes the dynamics of the various systems, including the orbital motions

of the weapons platforms and the trajectories of the targets and decoys in the gravitational field. The parameter λ is a "discount factor." The constraint sets X and U reflect physical or operational constraints. (There may be many other types of constraints. Our main concern at this point is an assessment of the computational problem, not the precise model structure.)

3.1.1 Computational Complexity

The optimal cost $V(s,x)$ for the abstract problem (1) is found by solving the Hamilton Jacobi equation of dynamic programming:

$$\min_{u\in U}\{b(x,u)\cdot\nabla V + c(x,u) + \Delta V - \lambda V\} = \frac{\partial V}{\partial s} \tag{2}$$

which also gives the optimal strategy for problem (1) in feedback form.

The numerical solution of system (1) is virtually impossible when the number N of state variables is large. The problem is not simply one of numerical analysis, but an irreducible difficulty. No matter what (numerical) approximation method is used, achieving a given level of precision as the dimension of the state space increases will require an exponentially increasing computational cost. This is not even a consequence of the optimization formulation – the associated linear eigenvalue problem has the same computational complexity. In effect, the computational problem is "NP–complete."

It is natural to attempt to avoid this problem by assigning individual

controllers (weapons platforms) to a portion of the state space, and to select control laws in an optimal fashion to deal with just that subset of the state space. This is optimization in the "class of local feedbacks." Unfortunately, in the absence of certain structural conditions on the system dynamics, this problem is *more demanding* computationally than the previous one.

To see that this is case, let I be an index set for the subsystems, $I = \{1, 2, \ldots, k\}$ and let n_i and m_i denote the number of states and controls respectively in subsystem $i \in I$. A local feedback is a mapping S_i from $[s, T] \times R^{n_i}$ into $U_i \subset R^{m_i}$, the set of admissible values of the control for subsystem i. Let $\mathcal{S}_L = \{S = (S_1, \ldots, S_k)\}$ be the class of local feedbacks. If u_S is a local feedback control and x_S is the corresponding solution, then optimization in the class $\mathcal{S}_L$ of local feedbacks is the problem

$$\begin{gathered} V(s,x) = \min_{S \in \mathcal{S}_L} \{E_S[\int_s^T e^{\lambda t} c(x_S(t), u_S(t)) dt | x_S(0) = x]\} \\ \dot{x}_S = b(x_S(t), u_S(t)) + \dot{w}(t) \qquad (3) \\ x_i(t) \in X_i \subset R^{n_i}, \quad u_i(t) \in U_i \subset R^{m_i} \end{gathered}$$

If we let $p^S(t,x)$ be the probability density of $x_S(t)$ corresponding to local feedback $S \in \mathcal{S}_L$, then a given strategy $R \in \mathcal{S}_L$ may be improved by the algorithm

Step 1: Compute p^R

Step 2: Solve

$$\begin{cases} \mathcal{L}_S V^S + c^S = 0, & V^S(T, \cdot) = 0 \\ S \in \text{Arg}\, \{\min_Z \{H(t, Z, p^R, V^S)\}, & S \in \mathcal{S}_L \} \end{cases} \qquad (4)$$

Here $c^S = c(x, S(x)), S \in \mathcal{S}_L$, $H(t, R, p, V)$ is the Hamiltonian of problem (3), and p^R is the solution of

$$\mathcal{L}_R p^R = 0, \quad p^R(0, \cdot) = \mu \tag{5}$$

$$\mathcal{L}_S = \frac{\partial}{\partial t} + \sum_j b_j^S \frac{\partial}{\partial x_j} + \sum_{i,j} a_{ij} \frac{\partial^2}{\partial x_i \partial x_j} \tag{6}$$

where μ is the initial density of the state $x(0)$, $\{a_{ij}\}$ is the covariance matrix of the noise, and $b^S(t, x) = b(t, x, S(t, x))$.

A fixed point of this algorithm satisfies the conditions of optimality for the problem (3). However, it is clear that the algorithm (4) requires more computations than simple dynamic programming in equation (2).

3.1.2 Systems with Decoupled Dynamics

If the underlying dynamics of the target system are decoupled, and if the controls (weapons) respond only to certain subsets of the state space, then team theoretic (local control) strategies can be computed. The SDI engagement scenario has this structure.

In this case we must have $b(t, x, u) = [b_1, \ldots, b_k]$ with

$$b_i(t, x, u) : [0, \infty) \times R^{n_i} \times U_i \to R^{n_i} \tag{7}$$

and we assume that the noises are not coupled between subsystems. Then, for each local control strategy $R = [R_i] \in \mathcal{S}_L$, the probability density of the

state x satisfies $p^R = \prod_{i \in I} p_i^{R_i}$, and it may be computed from equation (5) with the expression (7) substituted into the operator (6). Let $\mathcal{L}_{i,R_i}$ be the operator with the substitutions. The controls are still chosen through the combined performance index in problem (1). The functionals

$$c_i^R = \int c(t,x,R(x)) \prod_{j \neq i} p_j^{R_j}(t,x_j)dx_j, \qquad i \in I \tag{8}$$

are the conditional expectations of the instantaneous cost based only on knowledge at the i^{th} subsystem. Using these functions, a sufficient condition for a strategy S to be optimal agent by agent is

$$\min_{R_i} \left[\mathcal{L}_{i,R_i} V_i + c_i^R\right] = 0, \qquad i \in I \tag{9}$$

The corresponding optimal cost is $\mu_1(V_1) = \ldots = \mu_k(V_k)$ with

$$\mu_i(V_i) = \int_{R^{n_i}} \mu_i(dx_i)\, V_i(s,x_i) \tag{10}$$

where μ_i is the initial probability distribution of the subsystem state x_i.

This result provides an algorithm for the computation of a strategy optimal agent by agent.

Given $\epsilon, \nu \in [0,\infty)$:

Step 1: Choose $i \in I$, solve equation (9). If $\mu_i(V_i) \leq \nu - \epsilon$, then set $\nu := \mu_i(V_i)$ and

$$R_i := Arg \min_{R_k} [\mathcal{L}_{i,R_i} V_i + c_i^R] \tag{11}$$

If not, then choose another $i \in I$ until $\mu_i(V_i) \geq \nu - \epsilon, \forall i \in I$.

Step 2: When $\mu_i(V_i) \geq \nu - \epsilon, \forall i \in I$, then set $\epsilon := \epsilon/2$, and go to Step 1.

This algorithm produces a decreasing sequence $\nu^{(n)}$ which converges to a cost which is optimal agent by agent. A proof of convergence for the discrete version of this algorithm is given in [44]. (Even at this level of model simplification, it may be difficult to solve the problem in Step 1. We shall discuss two procedures for reducing this problem further in subsequent sections.)

The analysis in [44] establishes the same sequence of arguments – that is, a procedure for reducing the computational requirements of optimal stochastic control problems – for Markov chain – queuing system models of controlled discrete time systems. Such models may be useful in developing high level strategies for SDI interception operations. For example, simple birth and death type processes may be useful in describing the transition of targets through a region, particularly in cases where bursts of decoys are generated by a hard target during the transition.

These models, which do not account for the trajectories of the targets and decoys, may be useful in deciding the commitment levels of weapons within the region and in neighboring regions. Since they may be resolved by efficient "index rules" and "stochastic gradient" methods, they support "rapid prototyping development" of computational algorithms. In the next three sections we describe the development of efficient computational algorithms for models with special structures compatible with the mid-course phase of SDI operations framework.

3.1.3 Systems with Product Form Performance Measures

The assumption that the dynamics of the system be completely decoupled is unrealistic in most SDI operations scenarios. In the first instance enemy missiles will likely be launched in volleys, and groups launched from the same geographical site may travel together on route to a designated target. Alternately, targets may release a family of associated decoys during the course of a flight. The distributions (of the state vectors) of the decoys and the parent target will be dependent, at least for the initial portions of their flights. If, however, the distributions can be well approximated by independent distributions in the limit of large times, that is, as the distribution of decoys about the main target(s) stabilizes, then it is possible to compute agent by agent optimal engagement strategies using the second algorithm discussed above.

We shall omit most of the technical details, noting only the main points. First, it is necessary to have the probability distributions of the states of the target systems converge in the limit of "long times" to ergodic distributions with the product form

$$p(x) = c \prod_{i=1}^{k} p_i(x_i), i \in I \tag{12}$$

with c a normalization constant. By "long time" we mean times long relative to the time constants of the release process for decoys, for instance. If the release and "blooming" of the decoy configuration take place over a matter of a few minutes, then the total time of the mid-course phase may be considered long relative to the initial period of development. Second, the control problem corresponding to this situation is either the system (1) with $T = \infty$ or the

"ergodic control problem" with the average cost

$$\min_{S_L} \lim_{T\to\infty} \frac{1}{T} \int_0^T c(x(t), S(x(t))\, dt. \tag{13}$$

This class of diffusion process models may be regarded as the natural limits of Jackson networks of queues [29] under the scaling

$$x \to \frac{x}{N}, \quad t \to \frac{t}{N^2} \qquad \text{as } N \to \infty \tag{14}$$

As a description of an engagement between weapons and a target system, the queues correspond to the targets exposed to a given weapon system (the "server"). The output rate corresponds to the rate at which targets are passed on to neighboring weapons systems. The transition probabilities m_{ij} describe the likelihood of a target passing through weapons region i and entering weapons region j.

Efficient algorithms for the solution of control problems for networks of Jackson type were given in [44]. They have been applied to large scale systems including numbers of state variables (several hundred) that might be reasonably associated with subsystems of an SDI post-boost or mid-course engagement system. We shall describe algorithms for optimization of models of this type based on index rules in a subsequent section.

There are two ways in which the queuing network model may be associated with a differential equation model like the abstract system in problem (1). First, the queuing model can serve to provide a difference equation numerical approximation of the solution of the continuous time control problem. This is the approach developed in [33] which has become a standard

technique in the solution of control and scheduling problems. This method preserves the information on the trajectory and orbital dynamics contained in the differential equations. As we have argued, this information is important during the long period of the mid-course phase, when significant motions of the weapons platforms will take place.

Second, the diffusion process model in the model (1) may itself be an approximation to a queuing model in which there are a large number of elements, i.e., in the scaling in equation (14) with N the total number of targets/decoys in the model [31]. The use of diffusion approximations to represent large populations is a common technique; however, we are not aware of any studies which have determined that this would be an effective class of models for any phase of SDI operations. We have not pursued this point in this project; rather, we use queuing models as a component of the numerical analysis of the system model, including the required descriptions of the orbital mechanics and target features such as aspect angle.

3.2 Monte Carlo and Stochastic Gradient Methods

From the previous sections we have seen that it will be possible to compute the optimal local feedback controls which are the engagement strategies only under certain limited conditions. There may be circumstances when the information available for the design and execution of local strategies is poor. In this case we may have an *a priori* design for an engagement strategy, and it would be useful to have a method for evaluating and implementing

this strategy. One technique for accomplishing this is to parameterize the strategy, and optimize the parameter in an "open loop" mode using a Monte Carlo technique.

This is the underlying idea of the stochastic gradient method used in the theory of stochastic approximations [34, 42]. This method has been applied to advantage in large scale planning systems. It is easy to implement; it is efficient; and it can be readily adapted to treat optimization problems with integer valued variables. For example, it has been shown to be far more effective in producing network designs in specific instances than a very efficient simplex algorithm [13]. The primary use for this class of algorithms is in setting up a prototype weapons allocation system which can be systematically enhanced and upgraded in subsequent phases of the project.

Once again, we shall explain the method and the algorithms it implies in terms of an abstract dynamical optimization problem. (We discuss "static" problems a little later.) The method is general and can be applied to almost any class of optimization problems. See [34] for examples.

Consider the system

$$\min_u E[\int_0^T c(t, x(t), u(t))\, dt] \tag{15}$$

$$\dot{x}(t) = b(t, x(t), u(t)) + \dot{w}(t)$$

$$x \in R^N, u \in R^M$$

We make the feedback transformation

$$u(t) = S(t, x(t), v(t)), \quad v \in R^p \tag{16}$$

with the function $S : [0, \infty) \times R^N \times R^p \to R^M$ given. This is the *a priori* strategy. Since we wish to compute the best such strategy (parameterized by time functions $v(\cdot)$) in "open loop" form, we approximate the probability law P of the noise in terms of the empirical distribution

$$\mu = \frac{1}{r} \sum_{j=1}^{r} \delta_{\omega_j}(\omega) \tag{17}$$

where ω_j are trajectories of the noise obtained by a random generator and δ is the Dirac delta function. Now we must solve the deterministic optimization problem

$$\begin{aligned} \dot{x}^j(t) &= b(t, x^j(t), S(t, x(t), v(t))) + \dot{w}^j(t) \\ \min_v \frac{1}{r} \sum_{j=1}^{r} &\int_0^T c(t, x^j(t), S(t, x^j(t), v(t)))\, dt \end{aligned} \tag{18}$$

where $w^j(t)$ is a particular trajectory of the noise. This is a deterministic optimization problem, which we can solve by a gradient technique or the Pontryagin minimum principle.

The idea of the stochastic gradient method is similar to this, but it uses a recursive procedure to optimize the parameter $v(\cdot)$ of the strategy. Let $J(v)$ be the integral in expression (15). Suppose we are able to compute the gradient $\partial J/\partial v$ by an adjoint method (numerically, after a discretization in which $v(\cdot)$ is finite dimensional). Then the stochastic gradient algorithm is the recursive procedure

$$v_{r+1} = P_V[v_r - \rho_r \frac{\partial J}{\partial v}(v_r, \omega_r)], \quad \rho_r \in [0, \infty), \quad \forall r = 1, 2, \ldots \tag{19}$$

for any sequence of positive numbers $\{\rho_r\}$ with $\sum_r \rho_r = \infty$, and $\sum_r \rho_r^2 < \infty$. Here ω_r is a (numerically) generated realization of the process noise, and P_V is projection on the (finite dimensional) set V where $v(\cdot)$ takes its values.

Under convexity conditions it is possible to show that the algorithm converges globally [13]. Local convergence results are given in [34]. Under favorable, but not unrealistic smoothness and convexity conditions, it can be shown that the convergence rate of the algorithm is optimal.

There is an additional technical point which must be addressed. The SDI interception problem requires the treatment of integer valued (random) variables. It is clear that conventional algorithms for integer programming will not be effective for this problem. It is possible to design some heuristic algorithms based on the stochastic gradient method which show promise for the treatment of optimization on integer valued variables.

Here is a simple modification of the basic stochastic gradient algorithm which has been shown by [22] to work well for problems with integer valued variables. Suppose we have to solve the problem on N^M, the set of M-tuples of natural numbers:

$$\min_{x \in N^M} [Ef(x)] \tag{20}$$

Consider the following algorithm

$$x_{n+1} = x_n - a\frac{\partial f}{\partial x}([x_n], \omega_n), \quad a \in (0, \infty) \quad \text{fixed} \tag{21}$$

where $[x_n]$ is the integer nearest to x_n. Evidently, the sequence $[x_n]$ cannot

converge. Rather, it moves on some recurrent set of points. Suppose for simplicity these points belong to the hypercube $[0,1]^M$. Let p_i^1 denote the (relative) visit frequency of $[x_n]_i$ (the i^{th} component) to the point 1. Let p_i^0 be the visit frequency to 0. Then the solution $[x_n]^*$ at a given step is determined by the maximum frequencies

$$[x_n]^* = \begin{cases} 1 & \text{if } p_i^1 > p_i^0 \\ 0 & \text{if } p_i^1 \leq p_i^0 \end{cases} \tag{22}$$

This procedure may be improved slightly by ordering the elements of $[x_n]$ according to a scheme depending on the visit frequencies. Effective results for the network planning problem were obtained [22] for this algorithm in 3 steps (ordering the elements) and 3000 iterations, requiring 2 minutes of computer time. This compares favorably with the stochastic gradient algorithm for a continuous variable.

Other, more complex algorithms for treating integer valued, stochastic approximation problems using penalization and a modification of the probability law are given in [22]. However, the more complex algorithms do not improve the simple one (21) (22) in any significant way.

3.3 Index Rules and Efficient Implementations

To achieve the objective of designing weapons allocation algorithms which reflect the overall structure of the SDI BM operations system, it will be necessary to provide even more efficient allocation and coordination strategies. In

this subsection we shall discuss an analytical framework for this. It involves formulation of the weapons allocation problem as a scheduling problem, using a particularly compact set of feedback strategies called "index rules" to implement effective allocation schedules.

3.3.1 Stochastic Scheduling and Single Station Weapons Allocation

To see how the weapons system allocation and coördination problem may be formulated as a scheduling problem, consider a simplified version: one station with several weapons assets indexed by $j = 1, 2, \ldots, N$, with no precedence relations among these assets. Suppose that each threat's dynamics are described by a stationary Markov chain [32]. When a threat is engaged by resource j, the BM system (on a superior level) receives an immediate reward $R(t) = R_j(x_i(t))$ and its state changes to $x_i(t+1)$ according to its transition rule. The states of the threats not engaged remain unchanged. In this simplified problem we can think of the "state" of the threat platform as:

$$x_i(t) = \begin{cases} 0, & \text{threat is homing on target,} \\ 1, & \text{threat is not homing on target,} \end{cases} \tag{23}$$

The reward is a device used to represent the instantaneous significance of each threat to the Battle Manager and the cost of using weapon system j (e.g., probability of revealing position). It summarizes (in a very simplified way) strategic doctrine and the rationale of the Battle Manager. In this

simple version the states of all threats are observed and the problem is to schedule the order in which the threats are engaged to maximize the expected present value of the sequence of *immediate* rewards

$$E[\sum_{i=1}^{\infty} a^t R(t)] \tag{24}$$

where $0 < a < 1$ is a fixed discount factor. This is a resource allocation problem [41] called the multi–armed bandit problem [52].

3.3.2 Index Rules

In the basic version of the multi–armed bandit problem there are N independent resources (machines). Let $x_i(t)$ be the state of resource (machine) $i = 1, 2, \ldots, N$ at time $t = 1, 2, \ldots$ In the simplest version of this problem at each t one must operate exactly one machine. If machine i is selected, one gets an immediate reward $R(t) = R(x_i(t))$ and its state changes to $x_i(t+1)$ according to a stationary Markov transition rule; the states of the idle machines remain frozen, $x_i(t+1) = x_j(t), j \neq i$. The states of all machines are observed and the problem is to schedule the order in which the resources are operated to maximize the expected present value of the sequence of immediate rewards (24).

This problem was first formulated in the 1940's. The essential breakthrough came when Gittins and Jones [17] showed that to each resource (machine) i is attached an index which is a function only of its state, and

that the optimal policy operates the resource with the largest current index. This "index rule" is important because it converts the original N dimensional problem into N one dimensional ones. The index was subsequently shown to be [16, 18]

$$\nu_i(x_i) = \max_{\tau > 1} \frac{E[\sum_{i=1}^{\tau-1} a^t R_i(x_i(t)) | x_i = x_i]}{E[\sum_{i=1}^{\tau-1} a^t | x_i = x_i]} \tag{25}$$

where the maximization is over all stopping times $\tau > 1$. This is the dynamic allocation index (DAI), interpreted as the maximum expected reward per unit of discounted time.

A direct, formal solution and interpretation can be given to the weapons system allocation and coordination problem in this framework. More important is the fact that indices can be computed efficiently and quickly given models for the threat dynamics (e.g., Markov transition models).

In our approach to the BM weapons allocation and coordination problem for the single platform problem each resource (machine) $i = 1, 2, \ldots, N$ is characterized by the pair of sequences

$$\{X^i(s), F^i(s)\}, \quad s = 1, 2, \ldots \tag{26}$$

$X^i(s)$ is the random reward obtained when i is operated for the s^{th} time and $F^i(s)$ is the information (a σ-field) about machine i gathered after it has been operated $(s-1)$ times. At each time exactly one machine must be operated. Thus, $t = t^1 + t^2 + t^3 + \cdots + t^n$ where $t^i = t^i(t)$ is the number of

times i is operated during $1, 2, \ldots, t$. The decision at time $t+1$ is based on the available information

$$F(t) = \vee_i F^i(t^i + 1), \;\; t = 1, 2, \ldots \tag{27}$$

A policy π is any sequence of decisions that satisfies this information constraint. The problem is then to find the policy π that maximizes

$$V(\pi) = E[\sum_{t=1}^{\infty} a^t X^{i(t)}(t^{i(t)}(t))|F(1)] \tag{28}$$

where $i(t)$ is the machine operated at time t.

In this general situation the index for resource (machine) i after it has been operated $(s-1)$ times is

$$\nu_i(s) = \max_{\tau > s} \frac{E[\sum_s^{\tau-1} a^t X^i(t)|F^i(s)]}{E[\sum_s^{\tau-1} a^t|F^i(s)]} \tag{29}$$

where the maximization is over all stopping times $s < \tau \leq \infty$ of $F^i(\cdot)$. The index rule is to operate the machine with the largest index.

Several extensions of the preceding framework are necessary to capture a realistic weapons engagement scenario: More than one weapons technique can be operated at a time, additional constraints due to precedence rules may appear, the size of the problem may be very large if parametric dependence is to be investigated, pre-emptive strategies must be considered.

3.4 Multiple Station Weapons Coordination

This is a much more difficult problem than the single platform case. It is a version of the "multi-server" scheduling problem which encompasses all the difficulties of multi-agent stochastic control. One must solve a large-scale dynamic programming problem which is intractable in general cases. For the purposes of this application demonstration, we shall simplify the problem by adopting a specific, suboptimal form for the solution. The presumed structure is a two-layer one with several coordinators (BM stations) on the top level and individual agents (computers) controlling separate stations/platforms, each equipped with one or more weapons on the lower level. The agents respond to commands from designated BM stations. Each agent uses a "local feedback strategy," deploying weaponss in response to his perception of the threat (perhaps as defined by the BM station) while observing operational and tactical constraints imposed by the BM station. The individual agents optimize their performance measures using an index rule, as discussed in the last section. The BM station commands the agents' activities under his command by imposing operational constraints to satisfy global operational objectives, including:

(i) Concealing the positions of units;

(ii) Establishing a priority for threat engagement;

(iii) Establishing precedence constraints for weapons deployment; and

(iv) Coordinating weapons operations with other tactical operations (EW deployment, maneuver control, etc.)

The individual agents communicate their actions to their BM station, providing the instantaneous state of each weapon system (a conditional probability), the index values of each weapon the agent controls, and the value of the agent's local performance measure. They may also communicate their perception of the threat if it differs from that provided by the BM station. The BM station communicates operational constraints, e.g., precedence constraints on weapons deployment among agents, threat information, and performance constraints to the agents. The latter are the Lagrange multipliers, sometimes called "coordination variables," in systems theory.

This formulation limits inter-agent communications by channeling them all through the upper level BM stations, which simplifies the problem tremendously. It allows different kinds of agents on the lower level, including automated stations and subordinate systems. It allows each agent to have different information about the threat, derived from local sensor facilities. By permitting each agent to use a locally optimal strategy subject only to constraints imposed by the BM station, computation of a near optimal strategy is reduced to a manageable level. The simplified structure also makes the development of simulation scenarios straightforward. Permitting agents to exercise local control strategies as discussed in earlier sections, enhances the survivability of the overall system. If communications to a command center (coordinator) are lost, the agents lose constraint specification updates, sensor

updates, and performance multipliers supplied by the weapons commander. However, they can continue to function, assuming they retain some access to the sensor data, by optimizing and executing their local control actions.

It is difficult to prescribe a precise optimal decision making structure for reconfiguring the system under stress; however, one can provide an expert system, e.g., a production rule system, which would both assist the local area coordination of weapons platform operations and guide the reconfiguration of the weapons C^2 system in times when some units or command centers are off-line. We have not addressed this problem in this project.

3.5 Sensor Scheduling

In this section we consider the problem of scheduling a suite of sensors for the optimal detection of targets. The *sensor scheduling problem*[6] involves the simultaneous selection of of a signal processing scheme (according to some performance measure) *together* with the subset of sensors that collect the data. The scheduling problem and the model on which it is based serves to illustrate the key ideas in the treatment of other scheduling problems based on models using stochastic differential equations. As such it is a "generic" example.

Applications of this concept include multiple sensor platforms and dis-

[6]The work in this section was supported in part by the Army Research Office under contract DAAG-39-83-C-0028. The results discussed in this section are based on [8, 7].

tributed sensor networks. On a platform with multiple sensors there is a need to coordinate the data obtained from the different sensors, which may include radar, infrared, and other sensor technologies. The data obtained from different sensors will likely be of varying quality (as a function of range, aspect, ambient noise, etc.), and systematic procedures are required for apportioning confidence to different data sets and for basing decisions on the composite data set. For example, radar trackers are more accurate at long range than are infrared trackers; the reverse is true at short range.

In sensor networks one needs to coordinate data collected from a large number of sensors distributed over a large geographical area. Conflicts must be resolved and a preferred set of sensors selected (on a given time interval) and utilized in detection, estimation, and/or control decisions.

Sensor scheduling should be carried out on the basis of optimizing reasonably defined performance measures. These should include not only terms allocating penalties for errors in signal processing (detection and/or estimation); but also, they should include costs for managing the sensor network – e.g., costs for (de)activating sensors, and for switching from one set of sensors to another. For example, activating a radar sensor on a platform increases the detectibility of the platform, and this should be accounted as a switching cost. Using a more accurate sensor with a more complete data output may entail higher bandwidth communications and the allocation of more computational power to that sensor. In certain networks use of a sensor may involve physical movement of that unit, and this incurs a cost.

3.5.1 A Model Problem

Consider the problem of estimating the signal process $x(t) \in R^n$ based on a collection of measurements $\{y^i(t), i = 1, 2, \ldots, M\}$ from sensors indexed by $i \in [1, \ldots, M]$. (Each $y^i(t)$ can be vector-valued.) Suppose $x(t)$ is defined by the diffusion process

$$dx(t) = f(x(t))dt + g(x(t))dw(t), \quad x(0) = \xi, \quad 0 \leq t \leq T \tag{30}$$

with values $x(t) \in R^n$. Suppose the measurements satisfy

$$dy^i(t) = h^i(x(t))dt + R_i^{\frac{1}{2}} dv^i(t), \quad y_i(0) = 0, \quad i = 1, 2, \ldots, M \tag{31}$$

with values in R^{d_i}. Here $w(\cdot), v^i(\cdot)$ are independent, standard, Wiener processes in R^n, R^{d_i}, respectively, and $R_i = R_i^T > 0$ are positive-definite, $d_i \times d_i$ matrices.

If we are given the set of measurements $\{y^i(s), s \leq t, i = 1, \ldots, M\}$, then the problem of (detecting) estimating $x(t)$ is a standard problem in nonlinear filtering theory [39]. Suppose, in contrast, that we may select among the various signals $y^i(\cdot)$ during certain intervals of time, and base our estimates of $x(t)$ on the best selection, which varies as a function of time. That is, we wish to determine the *optimal utilization schedule* for the suite of sensors, based on "running costs" for using sensors and "switching costs" for changing the set of active sensors.

Let $c_i(x)$ be the cost of using sensor i when the state of the signal is x, and let $k_{oi}(x)$ and $k_{io}(x)$ be the respective costs of turning off and turning

on the i^{th} sensor. The *signal processing objective* is to compute, at time T, an estimate $\hat{\phi}(T)$ of a given function $\phi(x(T))$ of the state. It is natural to use the least squares estimation error as the performance measure for this function

$$\text{Estimation Error } = E\{|\phi(x(T)) - \hat{\phi}(T))|^2\} \tag{32}$$

Now consider the problem of scheduling the sensors. First, it is necessary to define a *configuration* of sensors. Let $\mathcal{N}$ be the set of all possible sensor configurations. An element ν of $\mathcal{N}$ is an M-tuple of 1's and 0's. A 1 in position j means that the j^{th} sensor is *on*, a 0 means that the sensor is *off*. There are $N = 2^M$ elements in $\mathcal{N}$. A *sensor schedule* is a piecewise constant map $u(\cdot) : [0,T] \to \mathcal{N}$. Let $\tau_j \in [0,T]$ be the *switching times* for the sensor schedule u, that is, the time instants at which individual sensors are turned on or off. Let ν, ν' be the sensor configurations before and after a switching. Then the *cost of switching* is

$$k_{\nu\nu'}(x) = \sum_{\{i\in\nu\}\{i\notin\nu'\}} k_{io}(x) + \sum_{\{j\notin\nu\}\{j\in\nu'\}} k_{oj}(x) \tag{33}$$

The *total running cost* associated with a configuration $\nu \in \mathcal{N}$ is

$$c_\nu(x) := \sum_{\{j\in\nu\}} c_j(x) \tag{34}$$

In (33)(34) the symbol $\{i \in \nu\}$ denotes the indices of the entries in the vector ν occupied by a 1; i.e., the sensors which are "on." The symbol $\{i \notin \nu\}$ denotes the set of indices corresponding to sensors that are "off." We shall *assume* that the running and switching cost functions c, k are bounded and

continuous as functions of x. Moreover, we shall *assume* that the switching costs are bounded from below by a positive constant.

Using this notation, for each sensor schedule $u(\cdot) : [0,T] \rightarrow \mathcal{N}$, the measurements available during $[0,T]$ are

$$dy(t, u(t)) := h(x(t), u(t)dt + r(u(t))dv(t) \tag{35}$$

Here for $x \in R^n, \nu \in \mathcal{N}$

$$h(x,\nu) := \begin{bmatrix} h^1(x)\chi_{\{\nu\}}(1) \\ \vdots \\ h^i(x)\chi_{\{\nu\}}(i) \\ \vdots \\ h^M(x)\chi_{\{\nu\}}(M) \end{bmatrix} \tag{36}$$

a block column vector, where

$$\chi_{\{\nu\}}(i) := \begin{cases} 1, & \text{if the } i^{th} \text{ position in } \nu \text{ is } 1 \\ 0, & \text{otherwise} \end{cases} \tag{37}$$

Similarly, for $\nu \in \mathcal{N}$,

$$r(\nu) = \text{ Block Diagonal } \{R_i^{\frac{1}{2}}\chi_{\{\nu\}}(i)\} \tag{38}$$

Also, $v(t) = [v^1(t), v^2(t), \ldots, v^M(t)]^T$ is the compound observation noise process.

In this framework a sensor scheduling strategy is defined by an increasing sequence of switching times $\tau_j \in [0,T]$ and the corresponding sequence $\nu_j \in \mathcal{N}$ of active sensor configurations. Let

$$u(t) = \nu_j, \quad t \in [\tau_j, \tau_{j+1}); \quad j = 1, 2, \ldots$$

be the notation for a strategy.

We are interested in finding the optimal sensor scheduling strategy simultaneously determining the optimal estimator $\hat{\phi}$ for each active sensor configuration. Given a strategy and the associated estimator $\hat{\phi}$, the corresponding cost is

$$J(u(\cdot), \hat{\phi}) = E\left\{|\phi(x(T)) - \hat{\phi}(T)|^2 \right. \tag{39}$$

$$\left. + \int_0^T c(x(t), u(t))dt + \sum_j k(x(t), u(\tau_{j-1}), u_(\tau_j)) \right\}$$

where we have introduced the notation

$$c(x, \nu) = c_\nu(x), \quad x \in R^n, \nu \in \mathcal{N}$$

$$k(x; \nu, \nu') = k_{\nu,\nu'}(x), \quad x \in R^n, \nu, \nu' \in \mathcal{N}$$

The optimal scheduling/estimation problem is to find among all admissible scheduling strategies and associated estimators the pair achieving

$$\inf_{u(\cdot),\hat{\phi}} J(u(\cdot), \hat{\phi}) \tag{40}$$

3.5.2 A Stochastic Control Formulation

As shown in [7, 8], the optimization problem can be reformulated as an optimal stochastic control problem with "impulse type" controls. To simplify the problem, suppose $\phi(x) = x$. Then the optimal estimator (for any sensor configuration) is the conditional mean

$$\hat{\phi}(T) = E^{u(\cdot)}\{x(T)|\mathcal{F}_T^y\}$$

where $E^{u(\cdot)}$ is conditional expectation with respect to the probability distribution induced by $u(\cdot)$ and

$$\mathcal{F}_T^y = \sigma\{y(t, u(\cdot)), t \leq T\}$$

is the (σ–algebra of) measurements available from the scheduling strategy over the observation interval. Let $\mu(u, t)$ be the conditional probability measure of $x(t)$ given $\mathcal{F}_t^y$ on R^n. Then the conditional mean as a best estimate can be written as

$$\hat{\phi}(T) = \Phi(\mu(u, T)) = \int_{R^n} x d\mu(u, T) \tag{41}$$

which we regard as a vector–valued functional of $\mu(u, T)$.

As a result of this simple transformation the scheduling/estimation cost may be rewritten as a function of the scheduling strategy $u(\cdot)$ alone

$$J(u(\cdot)) = E^{u(\cdot)}\Bigg\{\| x(T) - \Phi(\mu(u, T)) \|^2 + \int_0^T c(x(t), u(t))dt \tag{42}$$

$$+ \sum_{j=1}^{\infty} k(x(\tau_j), u(\tau_{j-1}), u(\tau_j))\chi_{\tau_j < T}\Bigg\}$$

where $\chi_{\tau_j < T}$ is the characteristic function of the set of (random) events with $\tau_j < T$.

Since we assumed that the switching costs are bounded from below, if the observation interval is finite, then the optimal cost will be finite, and there will be only a finite number of switchings among the sensor configurations during $[0, T]$. Because the control for (42) is a pure switching control, we shall follow standard terminology and call it an *impulsive control* [9].

The optimization problem becomes the following impulse control problem: *Find an admissible impulsive control* $u^*(\cdot)$ *such that*

$$J(u^*(\cdot)) = \inf_{u(\cdot)\in U_{ad}} J(u(\cdot))$$

where U_{ad} is the set of all impulsive control laws adapted to the observations $\mathcal{F}^{y(\cdot,u(\cdot))}$.

This problem falls into the class of optimal stochastic (impulse) control problems with *partially observations.* It can be converted into a problem with complete observations by introducting an evolution equation – that is, a *Zakai equation* – for the conditional probability distribution of $x(t)$ based on the observations.

Let $p(t, u(\cdot))$ be the conditional probability measure

$$p(u(\cdot), t)(\phi) = E\{\xi(t)\phi(x(t))|\mathcal{F}^{y(\cdot,u(\cdot))}\}$$

for each control $u(t)$. Here

$$\xi(t) = \exp\left\{\int_0^t \tilde{h}(x(s), u(s))^T dz(s) - \frac{1}{2}\int_0^t \| \tilde{h}(x(s), u(s)) \|^2 dz(s)\right\}$$

defines the change of measure in the Girsanov transformation

$$\frac{dP^{u(\cdot)}}{dP}|_{\mathcal{F}_t} = \xi(t)$$

so that under the probability measure $P^{u(\cdot)}$ the process

$$v(t) = z(t) - \int_0^t \tilde{h}(x(s), u(s))ds$$

is a standard Wiener process.[7] The function $\tilde{h}$ is the vector in (36) with each element multiplied by $R_i^{-\frac{1}{2}}, i = 1, \ldots, M$.

For each control $u(\cdot)$, $p(u(\cdot),t)(\phi)$ is the *unnormalized conditional probability measure* of $x(t)$ given the observations $\mathcal{F}^{y(\cdot,u(\cdot))}$. This function is the "state vector" in the sensor scheduling problem. It satisfies

$$dp(u(\cdot),t) = L^* p(u(\cdot),t)dt + \delta(\cdot, u(t))^T dy(t, u(\cdot)) \tag{43}$$

$$p(u(\cdot),t) = p_0$$

where $y((t, u(\cdot))$ is the control (schedule) dependent observations process and

$$\delta(x,\nu) := \begin{bmatrix} R_1^{-1} h^1(x) \chi_{\{\nu\}}(1) \\ \vdots \\ R_i^{-1} h^i(x) \chi_{\{\nu\}}(i) \\ \vdots \\ R_M^{-1} h^M(x) \chi_{\{\nu\}}(M) \end{bmatrix}$$

Thus, the infinite dimensional quantity $p((u(\cdot),\cdot)$ becomes the state vector in the fully observed version of the problem. Using p we can write the estimation cost functional as

$$J(u(\cdot)) = E\left\{\Psi(p(u(\cdot),T) + \int_0^T < p(u(\cdot),t), C(u(t)) > dt \right. \tag{44}$$

$$\left. + \sum_{i=1}^{\infty} \chi_{\tau_i < T} < p(u(\cdot),\tau_i), K(u_{i-1},u_i) > \right\}$$

[7]The process $x(\cdot)$ retains its probability law under $P^{u(\cdot)}$ due to the independence of the noises and the initial conditions.

where

$$C(u_i) = c_{u_i}, \quad u_i \in \{1, 2, \ldots, N\}$$

$$K(u_i, U_j) = k_{u_i, u_j}(\cdot), \quad u_i, u_j \in \{1, 2, \ldots, N\}$$

and Ψ is a functional defined on measures[8]

$$E^u\{\| x(T) - \Phi(\mu(u,T)) \|^2\} = E\{\Psi(p(u(\cdot), T))\}$$

To summarize, this formulation converts the optimal sensor scheduling problem based on partial information – the noisy measurements y – which has a finite dimensional state space, into a problem with full state observations, but an infinite dimensional state space.

3.5.3 Solution of the Optimization Problem

A solution to the optimal sensor scheduling problem, specifically, a set of variational inequalities defining the transitions in the sensor configuration and the switching times, can be derived from a dynamic programming argument. Let $u(t) = j$ be a fixed sensor schedule, and let p_j be the corresponding density $p(\cdot, j)$. Then

$$dp_j = L^* p_j dt + p_j (\bar{h}^j)^T dz(t) \tag{45}$$

$$p_j(0) = \pi, \quad j \in \{1, 2, \ldots, N\}$$

[8]In fact, $\Psi(\mu) = \mu(\chi^2) \| \mu(\chi) \|^2 / \mu(1)$ where $\chi^2(x) = \| x \|^2, x \in R^n$, and μ is any finite measure on R^n such that $\mu(\chi), \mu(\chi^2)$ are defined.

Let $p_{j,\pi}$ denote the solution to (45). Set

$$\Phi_j(t)(F)(\pi) = E\{F(p_{j,\pi}(t))\} \tag{46}$$

Then Φ_j is a semigroup (because $p_j(t)$ is a Markov process) which we shall use to define the evolution of the cost in the scheduling process.

To simpilfy the presentation, consider the case $N = 2$. Let

$$C_i := C(i, \cdot), \quad i = 1, 2$$

$$K_1 := K(1,2)$$

$$K_2 := K(2,1)$$

and let $C_1(\pi) =< C_1, \pi >$ with the other quantities similarly defined.

Now consider the set of functionals $U_1(\pi, t), U_2(\pi, t)$ such that

$$U_1(\pi, t) \geq 0, \quad U_2(\pi, t) \geq 0$$

$$U_1(\pi, T) = U_2(\pi, T) = \Psi(\pi)$$

$$U_1(\pi, t) \leq \Phi_1(s - t) U_1(\pi, s) + \int_t^s \Phi_1(\lambda - t) C_1(\pi) d\lambda \tag{47}$$

$$U_2(\pi, t) \leq \Phi_2(s - t) U_2(\pi, s) + \int_t^s \Phi_2(\lambda - t) C_2(\pi) d\lambda \tag{48}$$

$$\forall s \geq t$$

and

$$U_1(\pi, t) \leq K_1(\pi) + U_2(\pi, t) \tag{49}$$

$$U_2(\pi, t) \leq K_2(\pi) + U_1(\pi, t) \tag{50}$$

These expressions have the following interpretation:

$$U_i(\pi, 0) = \inf_{\substack{u(0)=i \\ p(0)=\pi}} J[u(\cdot)], \quad i = 1, 2 \tag{51}$$

is the optimal sensor scheduling and estimation performance for the system starting at time zero with the initial configuration indicated. Suppose we start with $u(0) = 1$, then so long as (49) holds with strict inequality we should use schedule $j = 1$, since the optimal performance in this configuration is less than the cost of switching to configuration $j = 2$ and then continuing optimally thereafter - defined by the right side in (49). The optimal performance during the period prior to a switch is determined by (47), which holds with equality prior to the switch. The latter is the equation of dynamic programming which governs the choice of the optimal esimation law while sensor configuration $j = 1$ is being used.

At any time t when condition (49) holds with equality, then it is optimal to switch from configuration $j = 1$ to $j = 2$, and continue optimally thereafter. The sensor schedule is determined by the sequence of switching times. For example, suppose $i = 1$ in (51) and let

$$\tau_1^* = \inf_{t \leq T} \{U_1(p_1(t), t) = K_1(p_1(t)) + U_2(p_1(t), t)\} \tag{52}$$

Then τ_1 is the *optimal time* to switch from configuration $j = 1$ to $j = 2$. Let

$$p^*(t) = p_1(t), \quad t \in [0, \tau_1^*].$$

Next define

$$\tau_2^* = \inf_{\tau_1^* \leq t \leq T} \{U_2(p_2(t), t) = K_2(p_2(t)) + U_1(p_2(t), t)\} \tag{53}$$

Then τ_2^* is the *optimal time* to switch back from configuration $j = 2$ to $j = 1$. Let

$$p^*(t) = p_2(t), \quad t \in [\tau_1^*, \tau_2^*].$$

In this way the sequence of optimal switching times is constructed.

In the general case when there are $N \geq 1$ sensor configurations, then the computation of the switching times is based on the inequality

$$U_i(\pi, t) \leq \min_{\substack{j \neq i \\ j=1,\ldots,N}} \{K_{ij}(\pi) + U_j(\pi, t)\}$$

The system (47)-(50), appropriately modified constitutes a set of *quasi-variational inequalities* defining the optimal sensor scheduling problem.

3.5.4 Implementation of the Algorithm

The numerical treatment of systems of Quasi-Variational Inequalities has only just recently been attempted. The basic ideas are not substantially different from the treatment of the nonlinear partial differential equations – the Hamilton–Jacobi equation – of dynamic programming. There is one substantial difficulity, however. That is that the boundary of the domain on which the solution is defined – the optimal continuation policy between switchings depends on the solution. Specifically, the switching set is defined by the solution to the continuation condition.

The numerical treatment of optimal scheduling conditions was beyond the scope of this effort.

4 Preliminary Considerations: Decision Support Systems for Weapons Management

The framework we have defined for the problem of managing weapons resources deployed from different platforms, including a coordinator and the exchange of summary information by the local agents through the coordinator, effectively supports "heuristic optimization" stategies. This is important, since there is no hope of solving the multi–station coordination problem using conventional analytical methods.

4.1 Constraint Directed Reasoning

Consider, for example, the method of *constraint directed search* developed by M. Fox [14, 15] (see also [40]). The case study treated by Fox is job shop scheduling which involves the selection of a set of operations whose execution leads to the completion of an order; and the assignment of start and finish times and resources to each operation. The number of possible schedules grows exponentially with the number of orders, alternative production plans, the number of substitutable resources, and other parameters of the system. By fully integrating the constraints into the search/scheduling process it is possible to bound the generation and focus the selection of aternative solutions. In effect, this treats the job shop scheduling task by "constraint-directed reasoning" [14].

Fox defines a four "level" procedure for constraint-directed scheduling of orders:

Level 1 Selects an order to be completed based on prioritization rules, its category, and due date.

Level 2 Does a capacity analysis of the plant to determine the earliest start time and the latest finish time for each operation associated with the order. This determines time binding costraints which are effective at the next level.

Level 3 Does a detailed scheduling of all resources necessary to produce the order. A "beam search" method is used to select the schedule, based on a pre-search analysis examining the constraints associated with the order (determining the direction of the search), including a determination of whether any new constraints should be generated. Level 3 outputs reservation time bounds for each resource required for the operations in the chosen schedule.

Level 4 Selects the actual reservations for the resources which minimize the "work-in-process" time.

It is easy to draw certain parallels between this approach to selection of resource allocation schedules, and aspects of the BM weapons system C^2 problem. For example, the interaction of weapons deployed by the same platform and by neighboring platforms must be coordinated to avoid counter productive interference effects, as we have noted. This requires observing

causality and precedence relationships (among other variables). Constraint directed search procedures may be useful at some level in the C^2 system for the delineation of options (which satisfy all the constraints). The hierarchical structure of the algorithm is suggestive designing a "semi-automated" weapons management decision support system. For example, the kinds of tasks done on Levels 1 – 3 in Fox's system could be automated. The deployment decisions made on Level 4 would be the responsibility of the BM system based on the constraint information output by the algorithms at Level 3.

The system developed by Fox for job shop scheduling was not intended for applications like management and scheduling of weapons/EW resources and engagements. For example, it has no facilities to describe the adversary nature of SDI encounters and the attendant need to secure operations; it does not provide for continued service under stressed conditions (loss of units); it has no provision for evaluation and fusion of sensor data; etc. In addition it treats orders as isolated events. In SDI operations it is necessary to "track" the evolution of threats using a dynamic model of threats based on sensor information. Based this, SDI engagements involve dynamic allocation problems, as we have argued. The methodology of Fox has no (apparent) means for accomodating dynamical relationships among arriving orders. We have discussed Fox's work here here as an example of some of the good work now under way in AI applied to resource allocation problems and to illustrate the way heuristic methods developed in one AI application can be used to suggest treatments in other contexts.

4.2 Other Related Work

Other interesting work includes the BATTLE fire control system developed at NRL by Slagle and Hamburger [46], and the work on applications of AI to C^3I reported in [4, 6, 11]. The work in [12] is especially relevant to this project. Our treatment of the scheduling algorithm is more sophisticated than the branch and bound technique used in [12] (and the beam search method used in [14] for that matter). However, the modeling methodologies based on "flavors" used in [12] are very interesting.

From a more general point of view, the problem of coordinating BM operations over an extended theater should be addressed using "Distributed Artificial Intelligence" (DAI) methods. Some preliminary work on this, and additional references may be found in the report [1]. It is our judgment that the theory and methodology of DAI techniques is not well developed, and that an application of these techniques to tactical battle management including weapons operations is premature at this time. However, in the long run this may be an important area to pursue.

References

[1] J.M. Abram, et al., "Distributed Decision Making Environment," Interim report, Contract RADC-TR-82-310, Rome Air Development Center, December 1982.

[2] R. Agrawal, M.V. Hegde, and D. Teneketzis, "Asymptotically efficient adaptive allocation rules for the multiarmed bandit problem with switching costs," *IEEE Trans. Auto. Control,* vol. AC-33, 1988, pp. 899-906.

[3] V. Anantharam, P. Variaya, and J. Walrand, "Asymptotically efficient allocation rules for the multiarmed bandit problem with multiple plays - Part I: I.I.D. rewards, - Part II: Markovian rewards," *IEEE Trans. Auto. Control,* vol. Ac-32, 1987, pp. 968-976, pp. 977-982.

[4] S.J. Andriole, ed., *Applications in Artificial Intelligence,* Petrocelli, Princeton, 1985.

[5] M. Athans, P. hosein, and J. Walton, "Dynamic weapon–target assignment problems with vulnerable C^2 nodes," *Proc. 1988 Command and Control Research Symp.,* Monterey, 1988.

[6] A.J. Baciocco, "Artificial Intelligence and C^3I," in *Applications in Artificial Intelligence,* S.J. Andriole, ed., Petrocelli, Princeton, 1985, pp. 491-504.

[7] J.S. Baras and A. Bensoussan, "Optimal sensor scheduling in nonlinear diffusion processes," *Proc. IEEE Conf. Decision and Control,* Austin, 1988, pp. 2342-2345.

[8] J.S. Baras, G.L. Blankenship, and H.G. Kwatny, "Stochastic Stability, Control, and Scheduling," Final Report, Army Research Office, Contract No. DAAG-39-83-C-0028.

[9] A. Bensoussan and J.L. Lions, *Controle Impulsionnel et Inequations Quasi-Variationelles,* Dunod, Paris, 1982.

[10] D. Bertsekas. *Dynamic Programming and Stochastic Control,* Academic Press, New York, 1976.

[11] D.A. Brown and H.S. Goodman, "Artificial Intelligence Applied to C^3I," in Applications in Artificial Intelligence, S.J. Andriole, ed., Petrocelli, Princeton, 1985, pp. 505-519.

[12] J.F. Cunningham, et al., "Modelling an EW System Using an Object Oriented Approach," in *Proc. Expert Systems in Government Symposium,* K.N. Karmal, ed., IEEE Computer Society Press, Washington, 1985.

[13] R. Dodu et al, "Methodes de gradient stochastique pour l'optimization des investissements dans un reseau electrique," *EDF Bull.,* Serie C, 1981, pp. 133-164.

[14] M. Fox, "Constraint-Directed Search: A Case Study of Job-Shop Scheduling," Ph.D. Dissertation, Computer Science Dept., CMU, 1983.

[15] M. Fox, "Knowledge Representation for Decision Support," in *Knowldege Representation for Decision Support Systems,* L.B. Methlie and R.H. Sprague, eds., Elsevier, North-Holland, 1985, pp. 3-26.

[16] J. C. Gittins, "Bandit processes and dynamic allocation indices," *Journal of the Royal Statistical Society,* vol. 41, pp. 148-177, 1979.

[17] J. C. Gittins and D. M. Jones, "A dynamic allocation index for sequential design of experiments," in *Progress in Statistics, European Meeting of Statisticians 1972*, J. Gani, K. Sarkadi and I. Vince (Eds), vol. 1, pp. 241-266, North-Holland, 1974.

[18] J. C. Gittins and K. D. Glazebrook, "On Bayesian models in stochastic scheduling," *Journal of Applied Probability*, vol. 14, pp. 556-565, 1977.

[19] K. D. Glazebrook, "Scheduling tasks with exponential service times on parallel processors," *Journal of Applied Probability*, vol. 16, pp. 685-689, 1979.

[20] K. D. Glazebrook, "On randomized dynamic allocation indices for sequential design of experiments," *Journal of the Royal Statistical Society*, vol. 42, pp. 342-346, 1980.

[21] K.D. Glazebrook, "On stochastic scheduling with precedence relations and switching costs," *Journal of Applied Probability*, vol.17, pp. 1016-1024, 1980.

[22] M. Goursat et al, "Stochastic gradient methods for optimizing electrical transmission networks," preprint, 1985.

[23] J. Hwang, D. Schutzer, K. Shere and P. Vera, *Selected Analytical Concepts in Command and Control*, Gordon and Breach, 1982.

[24] H. Hamburger and J.R. Slagle, "Responding to a Call for Fire," *NRL Rept. 8768*, Dec. 1983.

[25] G.W. Hopple and S.M. Halprin, "Artificial Intelligence for C^3I: The Design and Development of a Prototype Causal Schema Knowledge-Based System," in *Proc. Expert Systems in Government Symposium,* K.N. Karmal, ed., IEEE Computer Society Press, Washington, 1985, pp. 160-168.

[26] K.N. Karmal, ed., *Proc. Expert Systems in Government Symposium,* IEEE Computer Society Press, Washington, 1985, pp. 406- 413.

[27] R.C. Kaste, "Concepts for a tactical fire control decision aid," in *Proc. Expert Systems in Government Symposium,* K.N. Karmal, ed., IEEE Computer Society Press, Washington, 1985, pp. 311-315.

[28] J.D. Kattar, "A solution of the multi–weapon, multi–target assignment problem," WP–26597, The MITRE Co., Bedford, February 1986.

[29] F.P. Kelly, Reversibility and Stochastic Networks, Wiley, New York, 1979.

[30] G. P. Klimov, "Time sharing service systems I," Theory of Probability and Applications, vol. 19, pp. 532-551, 1974.

[31] T.G. Kurtz, *Approximations of Population Processes,* SIAM CBMS-NSF Series, Philadelphia, 1981.

[32] H. Kushner, *Introduction to Stochastic Control,* Holt Rinehart and Winston, 1971.

[33] H.J. Kushner, *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations,* Academic Press, New York, 1977.

[34] H.J. Kushner and D. Clark, *Stochastic Approximation Methods for Constrained and Unconstrained Systems*, Springer-Verlag, Berlin, 1978.

[35] T.L. Lai, "Some thoughts on stochastic adaptive control," *Proc. 23rd IEEE Conf. Decision and Control*, Las Vegas, 1984, pp. 51-56.

[36] T.L. Lai and H. Robbins, "Asymptotically efficient allocation of treatments in sequential experiments," in *Design of Experiments*, T.J. Santner and A.C. Tambane, eds., Marcel–Decker, New York, 1984, pp. 127-142.

[37] T.L. Lai and H. Robbins, "Asymptotically efficient adaptive allocation rules," *Adv. Appl. Math.*, vol. 6, 1985, pp. 4-22.

[38] C.W. Li and G.L. Blankenship, "Optimal stochastic scheduling of systems with Poisson noises,"*Appl. Math. Optim.*, vol. 15, 1987, pp. 187-221.

[39] R.S. Lipster and A.N. Shiryayev, *Statistics of Random Processes*, Springer–Verlag, New York, 1977.

[40] P.S. Ow, "Heuristic Knowledge and Search for Scheduling," Ph. D. Dissertation, Computer Science Dept., CMU, 1984.

[41] M. Pinedo, "Stochastic scheduling with release dates and due dates," *Operations Research*, vol. 31, 1983.

[42] B.T. Polyak, "Convergence and convergence rate of iterative stochastic algorithms," *Automatica i Telemekhanika*, vol. 12, 1976, pp. 83-94.

[43] J.P. Quadrat, "On optimal stochastic control problems for large systems," preprint, 1982.

[44] J.P. Quadrat and M. Viot, "Product form and optimal local feedback for multi-index Markov chains," *Proc. Allerton Conf. Circuits and Systems,* Monticello, Ill. 1980.

[45] H. Robbins, "Some aspects of the sequential design of experiments," *Bull. Amer. Math. Soc.*, vol. 55, 1952, pp. 527-535.

[46] J.R. Slagle and H. Hamburger, "An Expert System for a Resource Allocation Problem," *Comm. ACM,* vol. 28, 1985, pp. 964-1004.

[47] J.R. Slagle, M.W. Gaynor, and H. Hamburger, "Decision Support System for Fire Support Command and Control," *NRL Rept. 8769,* (revised) July 1984.

[48] S.C.A. Thomopoulos and L. Zhang, "Distributed filtering with random smapling and delay," *Proc. IEEE Conf. Decision and Control,* Austin, 1988, pp. 2348-2353.

[49] J. Tsitsiklis, "A lemma on the multiarmed bandit problem," *IEEE Trans. Auto. Control,* vol. AC-31, 1986, pp. 576-577.

[50] P. Varaiya, J. Walrand and C. Buyukkoc, "Extensions of the multiarmed bandit problem: The discounted case," *IEEE Trans. Auto. Control,* vol. AC-30, 1985, pp. 426-439.

[51] G. Weiss, "Multiserver stochastic scheduling," in *Deterministic and Stochastic Scheduling,* M. A. H. Dempster et al. (eds.), D. Reidel, Dordrecht, Holland, 1982, pp. 157-179.

[52] P. Whittle, "Multi-armed bandits and the Gittins index," *Journal of the Royal Statistical Society,* vol 42, 1980, pp. 143-149.

[53] P. Whittle, *Optimization over TIme,* vol. 1, John Wiley, New York, 1982.

May 22, 1989

Air Force Office of Scientific Research
ATTN: LTC Nelson
AFOSR/NM
Bolling AFB, DC 20322

Contract No. F49620-87-C-0026

Dear LTC Nelson:

Please find enclosed the following regarding the above referenced subject contract.

1. Six (6) copies of the Final Report
2. Required Form DD 882, *Report of Inventions and Subcontracts*
3. Ten (10) copies of the final DD250 and an envelope addressed to DCASR Philadelphia.
4. An additional copy of the DD250 and a self addressed, stamped envelope

If the report is acceptable, please sign and forward the DD250s.

Please address technical questions to Dr. Gilmer Blankenship at the number above and questions regarding administrative matters to the undersigned at that same number.

Sincerely,

Amy C. Pfirrmann
Contracts Administrator